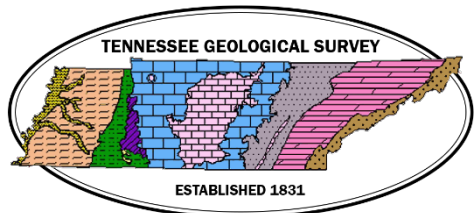# DIGITAL MAPPING TECHNIQUES 2024

The following was presented at DMT'24
May 13 - 16, 2024

The contents of this document are provisional

See Presentations and Proceedings
from the DMT Meetings (1997-2024)
http://ngmdb.usgs.gov/info/dmt/

Enhancing Map Database Production via Custom Scripts:

Adding Functionality to the GeMS Toolbox and Beyond

Andrew L. Wunderlich

DMT'24 Bloomington/Normal, IL

May 13-16, 2024

Department of
**Environment &
Conservation**

Division of Mineral and Geologic Resources

TENNESSEE GEOLOGICAL SURVEY

ESTABLISHED 1831

Developing Custom Scripts: Introduction

- Building geodatabases (GeMS or otherwise) to support geologic mapping and data preservation activities often requires the repeatable and consistent application of certain methods and/or calculations to features and classes within a geodatabase.

- We use the GeMS toolbox to accomplish tasks related to compliance, but there are many additional tasks that scripting can accomplish.

- In this talk:
  - New tools for GeMS-related work
  - Scripts for general geodatabase QC and cleanup
  - Scripts to work with graphics files (specifically PDFs)
  - The use of AI for script development
  - LIVE DEMOS OF A COUPLE SCRIPTS IF TIME PERMITS

Extensions to the GeMS Toolbox

- Tools designed to automate repetitive or cumbersome tasks common to many GeMS geodatabase creation/modification workflows

- Written in Python 3.x for use with ArcGIS Pro

- Use the existing GeMS scripting framework; take advantage of GeMS utility functions

- GeMS Toolbox (DMT'23):
  - Create and Edit:
    - **Attribute Points From Polys**
    - Migrate To New Database
  - Finalize:
    - **Remove Editor Tracking Fields**
    - **Translate to GeoPackage**

- GeMS Toolbox (DMT'24):
  - Create and Edit:
    - Get Glossary Terms
    - Get SourceIDs
    - Replace SourceIDs
  - Finalize:
    - Export Map Layers to .lyrx

Get Glossary Terms

- Written to help geologists prepare their databases BEFORE the internal GeMS Validation and QC process begins.
- Scans fields requiring definitions in a GeMS-y geodatabase:
  - Type
  - Any "Confidence" field (Identity, Existence, Scientific, GeoMaterial, etc.)
  - Any "Style" field (Paragraph, etc.)
  - Any "Units" field (Age, etc.)
- **Identifies classes/tables with <Null> values**
- Writes output to screen or to a text file
- Geologist can check against PARENT list and add missing terms or modify as necessary

Get SourceIDs and Replace SourceIDs

- Get SourceIDs: Help geologists prepare their databases BEFORE the internal GeMS Validation and QC process begins
  - Creates a list of all SourceIDs used in the database
  - **Identifies classes/tables with <Null> values**
  - Can ignore the DataSources table to prevent confusion and aid QC
  - Writes output to screen or to a text file

- Replace SourceIDs: Help make corrections during QC; revise older databases and metadata
  - Uses `key: value` pairs to find and replace SourceIDs in a database AND its accompanying standalone metadata xml/txt file
  - Robust and highly efficient find/replace capabilities thanks to use of regex
  - "Verbose results" option allows for detailed QC of tool output
  - MOST IMPORTANTLY:
    - Handles features with multiple IDs in one field; e.g., key: value pairs `A: 1` and `B: 2` will replace a field value "A|B" with "1|2" or "A|C|B" with "1|C|2"
    - Matches whole words only to prevent key substrings from replacing parts of other keys; e.g., pair `DAS001: ABC-123` won't erroneously find/replace "DAS001Z" with "ABC-123Z"

General-purpose scripts for map geodatabases

- Tools created to assist with the QC and general "hygiene" of geodatabases

- Written in Python 3.x for use with ArcGIS Pro

- Some have roots in GeMS framework but are applicable to any geodatabase

- QC Tools:
  - Describe (GeMS) Database
  - Recalculate Feature Class Extents

- Attribute Tools:
  - Attribute Lines From Points
  - Attribute Points From Points
  - Add and Calculate Lat-Lon Fields

QC Tools

- Recalculate Feature Class Extents:
  - Updates stored extent of feature classes which is not automatically done via editing or after geoprocessing
  - Uses `arcpy.management.RecalculateFeatureClassExtent()` method
  - Fixes issue with metadata tool reporting erroneous extent
  - Good hygiene for databases before publishing

- Describe (GeMS) Database:
  - Works on ANY geodatabase or GeoPackage (not only GeMS)
  - Relies on the **gdb_object_dict** routine of the GeMS_utilityFunctions to inventory the input database
  - For a feature class, displays the info of over 50 internal attributes
  - Produces a report in the output window, optionally writing the report to a text file

Describe (GeMS) Database

```
ContactsAndFaults:
    catalogPath: C:\GIS\PROJECTS\FY2023 NGGDPP Project\BraytonTN_GeologicMap\BraytonTN_GeologicMap.gdb\GeologicMap\ContactsAndFaults
    FIDSet: None
    aliasName: ContactsAndFaults
    areaFieldName:
    attributeRules: []
    baseName: ContactsAndFaults
    canVersion: False
    changeTracked: False
    children: []
    childrenExpanded: True
    createdAtFieldName:
    creatorFieldName:
    dataElementType: DEFeatureClass
    datasetType: FeatureClass
    dataType: FeatureClass
    defaultSubtypeCode: -1
    DSID: 59
    editedAtFieldName:
    editorFieldName:
    editorTrackingEnabled: False
    extension:
    extensionProperties: {}
    extent: 668049.518299997 115870.010400001 679789.9978 129811.3697
    featureType: Simple
    fields: ['OBJECTID', 'SHAPE', 'Symbol', 'Type', 'IsConcealed', 'LocationConfidenceMeters', 'ExistenceConfidence', 'IdentityConfidence', 'Label', 'DataSourceID', 'Notes',
            'ContactsAndFaults_ID', 'created_user', 'created_date', 'last_edited_user', 'last_edited_date', 'SHAPE_Length']
    file: ContactsAndFaults
    fullPropsRetrieved: True
    geometryStorage:
```

Describe (GeMS) Database

```
ContactsAndFaults (continued):
    globalIDFieldName:
    hasGlobalID: False
    hasM: False
    hasOID: True
    hasSpatialIndex: True
    hasZ: False
    indexes: ['FDO_OBJECTID', 'FDO_SHAPE']
    isCOGOEnabled: False
    isCompressed: False
    isTimeInUTC: True
    isVersioned: False
    lengthFieldName: SHAPE_Length
    metadataRetrieved: False
    MExtent: nan nan
    modelName:
    name: ContactsAndFaults
    OIDFieldName: OBJECTID
    path: C:\GIS\PROJECTS\FY2023 NGGDPP Project\BraytonTN_GeologicMap\BraytonTN_GeologicMap.gdb\GeologicMap
    rasterFieldName:
    relationshipClassNames: []
    representations: []
    shapeFieldName: SHAPE
    shapeType: Polyline
    spatialReference: WKID 32136; NAD_1983_StatePlane_Tennessee_FIPS_4100
    subtypeFieldName:
    versionedView:
    ZExtent: nan nan
    feature_dataset: GeologicMap
    concat_type: Simple Polyline Feature Class
    gems_equivalent: ContactsAndFaults
```
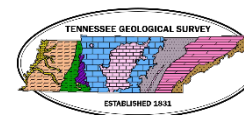
Attribute Tools

- Attribute Points/Lines From Points:
  - Works on ANY geodatabase
  - Uses an "in-memory" spatial join to transfer an attribute from a point class to an attribute of a point/line class which is nearest to it (within the search radius)
  - Tool inputs are filtered to prevent choosing incompatible feature types

- Add and Calculate Lat-Lon Fields:
  - Adds and calculates fields to point features
  - If fields already exist, simply calculates values
  - Allows user to choose precision of output
  - Uses `arcpy.management.CalculateGeometryAttributes()` method which respects the Geographic Transformations environment setting
  - Useful for calculating X, Y in a CS other than the native CS

Graphics Scripts

- Scripts run outside of ArcGIS, but could be added to a Python toolbox with minor modification
- Use **GraphicsMagick**: "the swiss [sic] army knife of image processing"
  - Collection of tools and libraries that support read/write and manipulation of around 100 file formats
  - Works with DPX, GIF, JPEG, JPEG-2000, JXL, PNG, PDF, PNM, TIFF, etc.
    - Convert an image from one format to another (e.g., TIFF to JPEG)
    - Resize, rotate, sharpen, color reduce, or add special effects to an image
    - Compare two images
    - Turn a group of images into a GIF animation sequence
    - Create a composite image by combining several separate images
    - Draw shapes or text on an image
    - Describe the format and characteristics of an image
  - Capable of creating images on the fly: great for web apps

Graphics Scripts

- GetPDFAttributes:
  - Gets page attributes of PDF files in a folder and its subfolders (recursive)
  - Returns a summary of the Width and Height (in inches) of pages in each document
  - Useful for finding oversize/different size pages in a large collection

- MakePDFThumbnails:
  - Creates thumbnail images of PDF files in a folder and its subfolders (recursive)
  - User can adjust the image size
  - Can automatically trim excess white space from edges
  - Useful for producing thumbnail images for creating a catalog, browse graphics for a website, etc.

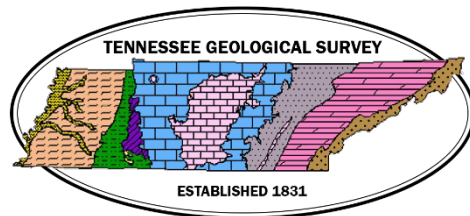# AI for help with scripting/automation

- Use ChatGPT to get help!
  - Ask a question, get Python code that works(ish)! Really(ish)!
  - Trick is to develop skills at **prompting** to return valid results:
    - Test the code
    - Challenge the answers you receive
    - Report feedback and you will get updates to the code based on your responses

Live demo if there's time!

Links to scripts and utilities

- GeMS GitHub Discussions page:
  - https://github.com/DOI-USGS/gems-tools-pro/discussions
- GraphicsMagick:
  - http://www.graphicsmagick.org/
- GitHub:
  - https://github.com/alwunder